DTIC FILE COPY

(2)

AD-A229 993

Progress in Characterizing Strictly Unidimensional IRT Representations

Brian W. Junker[1]

Department of Statistics
University of Illinois at Champaign-Urbana

November 29, 1990

DTIC
ELECTE
DEC 26 1990
S B D

Prepared for the Cognitive Science Research Program, Cognitive and Neural Sciences Division, Office of Naval Research, under contract number N00014-90-J-1984, 4421-560—01. Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[1]Current address: Department of Statistics, Carnegie Mellon University, Pittsburgh PA 15213.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION: Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION: University of Illinois, Department of Statistics; 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION: Cognitive Science Program, Office of Naval Research (Code 1142CS)
6c ADDRESS (City, State, and ZIP Code): 101 Illini Hall, 725 S. Wright Street, Champaign, IL 61820	7b ADDRESS (City, State, and ZIP Code): 800 N. Quincy Street, Arlington, VA 22217-5000
8a NAME OF FUNDING/SPONSORING ORGANIZATION; 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-90-J-1984
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61153N	RR04204	RR04204-01	4421-560

11 TITLE (Include Security Classification)

Progress in characterizing strictly unidimensional IRT representations.

12 PERSONAL AUTHOR(S)

Brian W. Junker (Dept. of Statistics, Carnegie Mellon University, Pittsburgh, PA 15213)

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final Report	FROM Apr 1990 TO Sep 1990	1990 November 29	35

16 SUPPLEMENTARY NOTATION

To be submitted for publication in Annals of Statistics.

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	See reverse

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

See reverse.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION
22a NAME OF RESPONSIBLE INDIVIDUAL: Dr. Charles E. Davis	22b TELEPHONE (Include Area Code): (703) 696-4046; 22c OFFICE SYMBOL: ONR-1142-CS

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

Abstract

Considerable attention has been paid to the development of nonparametric conditions on $P[\underline{X}_J = \underline{x}_J]$ that characterize a $d_L = 1$ (locally independent, monotone, unidimensional) latent variable representation for the binary items $\underline{X}_J = (X_1, \ldots, X_J)$. Holland and Rosenbaum (Holland 1981; Rosenbaum 1984; Holland and Rosenbaum 1986) focus on the conditional association (CA) of subtest scores under strict unidimensionality, and Stout (1987, 1990) treats the larger class of essentially unidimensional ($d_E = 1$) models, focusing on consistent estimation of Θ using proportion correct $\overline{X}_J$ in long tests. In the present paper we investigate the intersection of these two approaches, using Stout's principle that any reasonable set of items $\underline{X}_J$ can be embedded in an infinitely long sequence of items $\underline{X}$ of the same character.

We introduce three concepts which are helpful in the search for such a characterization. First, we consider only representations which are minimally *useful* in the sense that (1) the latent trait Θ can be consistently estimated from the item responses; (2) Θ is monotonically related to the test's "true score;" and (3) Θ is not constant in the examinee population. Second, we argue that the condition $\mathrm{Cov}\,(X_i, X_j | \overline{X}_J) \leq 0$ is a natural one to add to CA and $d_E = 1$ to ensure that the items are locally independent with respect to Θ. Third, we show that the monotonicity of the empirical ICC's $P[X_j = 1 | \overline{X}_J - X_j/J]$ is intimately related to ICC monotonicity: this "manifest monotonicity" must hold if $d_L = 1$ holds; and conversely it can be used to verify monotonicity of the usual ICC's $P_j(\theta)$ when $d_E = 1$ holds.

We obtain a nearly complete nonparametric characterization of *useful* $d_L = 1$ representations in terms of CA, $d_E = 1$, manifest monotonicity, and the above negative covariance condition. The negative covariance condition may not be strictly necessary for all $d_L = 1$ representations, but we show with a small simulation study that it probably does hold for $d_L = 1$ models often considered in practice: the Rasch, 2PL and 3PL models.

Key Words: strict unidimensionality, essential unidimensionality, useful models, conditional association, negative association, empirical item characteristic curves, monotonicity, simulation.

Contents

DTIC COPY INSPECTED 5

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification

By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

1 Introduction

Item response theory, IRT, is a modern attempt to model—and statistically analyze—examinee responses on standardized achievement or aptitude tests. IRT modeling and analysis, which occurs at the level of individual test questions—items—is greatly facilitated by the assumption of *unidimensionality*, i.e. that the latent trait "driving" the item responses is a one-dimensional, typically real-valued, random variable. Birnbaum (1968) and Lord (1980) provide complete accounts of traditional unidimensional IRT. In this paper we are concerned with a general characterization of (the distributions of) item response data for which traditional unidimensional IRT representations exist.

For our purposes, a test is simply a vector of J items, or equivalently J binary (0/1) item response variables,

$$\underline{X}_J = (X_1, X_2, \ldots, X_J),$$

representing the correctness of responses of a randomly-chosen examinee to the J test items. Let $\underline{x}_J$ represent an arbitrary fixed outcome of $\underline{X}_J$, i.e. a *response pattern*; an IRT model makes assumptions on the conditional distribution $P[\underline{X}_J = \underline{x}_J | \underline{\Theta} = \underline{\theta}]$ which impose restrictions on the marginal distribution $P[\underline{X}_J = \underline{x}_J]$ through the integral

$$P[\underline{X}_J = \underline{x}_J] = \int P[\underline{X}_J = \underline{x}_J | \underline{\Theta} = \underline{\theta}] \, dF(\underline{\theta}). \tag{1}$$

Here $F(\underline{\theta})$ is the sampling distribution of the latent trait or trait vector $\underline{\Theta} = (\Theta_1, \ldots, \Theta_d)$ in the examinee population under discussion; thus our point of view is similar to that of Cressie & Holland (1983).

We can estimate the marginal distribution $P[\underline{X}_J = \underline{x}_J]$ by examining the response data from an actual test administration. We will call this the *manifest structure* of the test, since it can be identified to arbitrary accuracy by increasing the manifest data: in this case, by increasing the number of examinees observed. On the other hand neither the marginal distribution $F(\underline{\theta})$ nor the conditional distribution $P[\underline{X}_J = \underline{x}_J | \underline{\Theta} = \underline{\theta}]$ is directly observable to us, in the sense that neither quantity is fully identifiable without also increasing the number of items[1]. These quantities determine the *latent structure* of the test.

[1] It is possible in principle to decide whether the Rasch model holds, without more items, using the special

The representation in (1) does not itself restrict the distribution of item responses $P[\underline{X}_J = \underline{x}_J]$ in any way. Standard IRT practice involves the imposition of additional conditions that make (1) a restrictive, and hence meaningful, representation. It then becomes a meaningful and important question to ask whether the model so proposed "fits" the observed response data.

The traditional IRT assumptions are that *local independence* holds,

$$\begin{aligned} P[\underline{X}_J = \underline{x}_J | \underline{\Theta} = \underline{\theta}] &= \textstyle\prod_{j=1}^{J} P[X_j = x_j | \underline{\Theta} = \underline{\theta}] \\ &= \textstyle\prod_{j=1}^{J} P_j(\underline{\theta})^{x_j} (1 - P_j(\underline{\theta}))^{1-x_j}, \end{aligned} \tag{LI}$$

and that *monotonicity* holds,

$$P_j(\underline{\theta}) \equiv P[X_j = 1 | \underline{\Theta} = \underline{\theta}] \textit{ coordinatewise nondecreasing in } \underline{\theta}, \ \forall\, j \tag{M}$$

in the sense that if $\theta_k^{(1)} \leq \theta_k^{(2)}$ for all $k = 1, 2, \ldots, d$ then $P_j(\underline{\theta}^{(1)}) \leq P_j(\underline{\theta}^{(2)})$. When Θ is one-dimensional, $P_j(\theta)$ will be called an *item characteristic curve*, ICC.

One additional assumption is needed to make (1) restrictive, namely that the *dimensionality* d of $\underline{\Theta}$ is much smaller than the test length J (see for example Holland and Rosenbaum, 1986), that is:

$$d \ll J. \tag{D}$$

(In the development that follows, this is formalized by requiring that d remain fixed as J grows.) The three assumptions, LI, M, and D form the foundation of item/test modeling in traditional IRT. Appendix A gives examples to show that if any of these three assumptions is completely omitted the resulting "model" will fit any distribution of binary data (hence making it scientifically meaningless). The least d for which the representation (1) holds and satisfies LI and M (and smoothness of the IRF's) we will denote d_L. We will refer to the case in which $d_L = 1$ as the *strictly unidimensional* case.

Various special cases of the strict unidimensionality assumptions have been investigated to see what properties they imply for the manifest distribution $P[\underline{X}_J = \underline{x}_J]$ through (1). Holland and Rosenbaum (Holland, 1981; Rosenbaum, 1984; Holland and Rosenbaum, 1986)

relationship of this model with log-linear models (c.f. Cressie and Holland, 1983; Tjur, 1982). However this does not appear to be possible for other models, nor for the general question of unidimensionality prior to parametric model selection which concerns us here.

have shown that when the general $d_L = 1$ assumptions hold, the items $\underline{X}_J$ must be *conditionally associated*; this shows that $d_L = 1$ is a restrictive and hence meaningful set of conditions. Cressie and Holland (1983) (se also Tjur, 1982), have characterized the Rasch model in terms of a suitably restricted log-linear model for $P[\underline{X}_J = \underline{x}_J]$. And de Finetti's Theorem in classical probability theory may be used to characterize an infinite sequence of items with identical item response functions by the property that the items themselves must be *exchangeable*.

Stout (1987; 1990) capitalizes on the good θ-estimation properties of the proportion correct score $\overline{X}_J = \frac{1}{J}\sum_1^J X_j$, when J is large and $d_L = 1$ holds, to produce a statistical test of latent-trait unidimensionality. Stout's statistical test is tailored to his *essential unidimensionality* condition ($d_E = 1$) which, in contrast to strict unidimensionality, allows there to be some minor dependencies among items as well as nonmonotonicities of individual ICC's.

Since we will be considering latent variable representations that are somewhat more general than the traditional $d_L = 1$ representation, it is worthwhile to ask what constitutes a "useful" unidimensional latent variable representation. In a nonparametric setting we propose that such a representation should satisfy the following definition. Note that what we mean by "useful" here relates primarily to connecting an examinee's item responses with an estimate of or inference about his/her latent trait score. For other purposes, other definitions might be appropriate.

Definition 1.1 *An IRT representation, in which LI may or may not hold, will be called* useful *if and only if the following principles are satisfied:*

U1 Θ can be estimated from the observed values of $X_1, X_2, \ldots, X_J$. At minimum this should mean that there are functions $t_J(x_1, \ldots, x_J)$ that consistently estimate Θ in the sense that

$$t_J(X_1, \ldots, X_J) \to \Theta$$

as the test length J grows. Moreover, consistent estimation should still be possible even though any fixed small group of items items $(Y_1, \ldots, Y_{J_0})$ in $\underline{X}$ is dropped from $\underline{X}$.

U2 Examinees with higher Θ values tend to score higher on the test. A very weak condition along these lines is simply the requirement that the average ICC $\overline{P}_J(\theta)$ be increasing in θ for each J; in other words

$$E[\overline{X}_J | \Theta = \theta] \text{ is increasing in } \theta.$$

U3 Θ is useful for categorizing examinees. In particular Θ should be able to take on at the very least two distinct values, each with positive probability.

These principles are implicit in traditional IRT work, and are easily justified on practical grounds:

First, Θ has little statistical value as a index of ability, achievement, aptitude, or other latent trait, if it cannot be estimated; hence **U1**. There is no hope that Θ can be estimated with high precision unless $J \to \infty$ (e.g. the survey by Fienberg, 1986), so **U1** represents, in some sense, a minimal estimation condition. The principle that estimation of Θ should not depend strongly on which particular items are used is central to what we mean by "latent trait."

Principle **U2** reflects the interpretation of Θ as a quantity of the latent trait, and of the test as an instrument for measuring that quantity. It will be seen below (Theorem 2.2) that **U2** also makes it easier for the representation to satisfy **U1**. When specific parametric models are constructed (recently, e.g., Jannarone, 1986; Sympson, 1987), there seems to be considerable latitude available in violating **U2** and still having a representation which satisfies **U1**; but for nonparametric purposes and for the purpose of interpreting Θ, **U2** is appropriate. Note that **U2** does not require individual ICC's to be monotone; this requirement is only made of the test characteristic curve.

Principle **U3** simply reflects the practical desire to use the test to diagnose, assess or otherwise categorize examinees and examinee populations. If there is no variation in Θ then there is no sensible way to use the test in this way. In terms of (1), **U3** asserts that the prior Θ distribution does not concentrate at a single θ value.

Except for the special cases of the Rasch model and de Finetti's theorem, no other characterizations of $d_L = 1$ in terms of features of the manifest structure $P[\underline{X}_J = \underline{x}_J]$ seem to be known. A general characterization is important for several reasons. First, it allows

us to better understand the structure, identifiability and meaning of the representation (1) under $d_L = 1$. For example, a consequence of our work here is a better understanding of how "far" each of the CA and $d_E = 1$ approaches are from the general $d_L = 1$ assumptions. Second, it suggests that fit tests of the general $d_L = 1$ representation are possible, without resorting to specific parametrizations of the ICC's, prior Θ distribution, etc. Hence we could distinguish between a lack of fit due to latent trait multidimensionality and a lack of fit due to a poor choice of parametrization.

In this paper we review the nonparametric approaches of Holland/Rosenbaum and Stout, and consider some new conditions which bring us closer to a general characterization of strict unidimensionality. Throughout this paper we embed the finite test $\underline{X}_J$ in an infinite sequence of similar items

$$\underline{X} = (X_1, X_2, \ldots).$$

LI and other traditional IRT properties extend in a natural way to the infinite item sequence $\underline{X}$ by requiring that they hold, in a consistent fashion, in every finite-length test $\underline{X}_J$ taken from $\underline{X}$.

In addition to the CA and $d_E = 1$ assumptions, we argue that it is natural to require $\mathrm{Cov}\,(X_i, X_j | \overline{X}_J) \leq 0$ to ensure that the items are locally independent with respect to Θ. Also, we show that the monotonicity of the empirical ICC's $P[X_j = 1 | \overline{X}_J - X_j/J]$ is intimately related to ICC monotonicity: this "manifest monotonicity" must hold if $d_L = 1$ holds; and conversely it can be used to verify monotonicity of the usual ICC's $P_j(\theta)$ when $d_E = 1$ holds. (The result that $d_L = 1$ implies manifest monotonicity has been discovered independently by I. Molenaar (priv. comm.) and is based on Grayson's (1988) monotone likelihood ratio result for proportion correct scores.)

Now consider representations in which expected values of the form $E[f(\underline{X}_J)|\Theta = \theta]$ are continuous, but LI may or may not hold. Within this class of "smooth" latent trait representations, we can characterize useful $d_L = 1$ as follows (a more formal statement of the result is given in Section 4).

Characterization of $d_L = 1$. *For any infinite sequence of binary items $\underline{X}$ and latent trait Θ, a useful $d_L = 1$ representation (1) holds, if and only if the following conditions hold:* CA.

$d_E = 1$, *manifest monotonicity, and*

$$\mathrm{Cov}\,(X_i, X_j | \overline{X}_J, \Theta) \leq 0, \ \forall\, i,j \leq J, \ \forall\, J. \tag{2}$$

Moreover in many practical settings it appears that the covariance in (2) continues to be negative if one omits the conditioning on Θ, because $\overline{X}_J$ is "nearly sufficient" for Θ. In the Rasch model for example, (2) is guaranteed to be negative when Θ is omitted. A small simulation study is reported in Appendix C, illustrating similar behavior in other logistic IRT models. Thus, although (2) spoils a characterization in terms of the manifest structure $P[\underline{X}_J = \underline{x}_J]$ alone, it appears we are quite close in practical situations.

2 Conditional association and essential independence

2.1 Conditional association

Holland and Rosenbaum have sought covariance conditions, or equivalently probability inequalities, in the distribution of $\underline{X}_J$ which must be satisfied if any $d_L = 1$ model applies. The starting place for their investigations may be taken to be coordinate-wise nondecreasing functions $f(\underline{y})$ of finite subtests $\underline{Y} = (Y_1, \ldots, Y_{J_0})$ taken from $\underline{X}$. Examples include

- The weighted average $f(\underline{Y}) = \frac{1}{J_0}\sum_1^{J_0} a_j Y_j$, $\forall\, a_j \geq 0$;
- The "all or nothing" score $f(\underline{Y}) = \prod_1^{J_0} Y_j$;
- The "at least one" score $f(\underline{Y}) = \max\{Y_j : j = 1, \ldots, J_0\}$;
- Item scores $f(\underline{Y}) = Y_j$.

The coordinatewise nondecreasing functions are exactly those scoring methods which assign more credit as examinees get more answers correct.

Under LI, any two such scoring methods will be positively correlated at each fixed ability level $\Theta = \theta$ (Rosenbaum, 1984): for all finite subtests $\underline{Y}$ taken from $\underline{X}$ and all coordinatewise nondecreasing functions $f(\underline{y})$ and $g(\underline{y})$,

$$\mathrm{Cov}\,(f(\underline{Y}), g(\underline{Y}) | \Theta = \theta) \geq 0, \tag{3}$$

for each possible θ. This condition can be converted into a condition on the latent structure into to a condition on the manifest structure:

Theorem 2.1 *(Rosenbaum, 1984; Holland and Rosenbaum, 1986). If $\underline{X}$ satisfies $d_L = 1$, then $\underline{X}$ is* conditionally associated (CA): *For every pair of disjoint, finite subtests $\underline{Y}$ and $\underline{Z}$ in $\underline{X}$, every pair of coordinatewise nondecreasing functions $f(\underline{Y})$ and $g(\underline{Y})$, and every function $h(\underline{Z})$,*

$$\text{Cov}\,(f(\underline{Y}), g(\underline{Y})|h(\underline{Z}) = c) \geq 0 \;\forall\; c \in \text{range}(h). \tag{CA}$$

Intuitively, a $d_L = 1$ test possesses so much internal coherence (the item responses are driven monotonically by the single latent variable Θ) that all reasonable subtest scores must be correlated, in any subpopulation of examinees selected by *any* criterion $h(\underline{z})$ relating to another part of the test.

Our statement of Theorem 2.1 is an easy extension of Holland and Rosenbaum's result for finite length tests to infinite item sequences. Note also that for finitely many binary (or indeed discrete) random variables $Z_1, Z_2, \ldots, Z_M$, conditioning on a scalar-valued function $h(\underline{Z})$ is equivalent to Holland and Rosenbaum's practice of conditioning on vector-valued $\underline{h}(\underline{Z})$. Seminal special cases of (3) and CA were developed by Holland (1981).

CA represents a wide variety of probability inequalities which can be tested in the manifest distribution $P[\underline{X}_J = \underline{x}_J]$. When CA fails the items cannot be treated as having a $d_L = 1$ latent representation. Applications to studying the internal coherence of a set of items may be found in Rosenbaum (1984) or Holland and Rosenbaum (1986). Related work appears in Holland (1981), Rosenbaum (1985), Rosenbaum (1987), and Rosenbaum (1988). An application of CA to assessing the dimensionality of standardized tests for the National Assessment of Educational Progress is described by Zwick (1987).

2.2 Essential independence

A successful approach to identifying unidimensional latent structure outside the strict $d_L = 1$ framework has been pursued in the seminal work of Stout (1997, 1990), and extended by Junker (1988, 1991). The main idea, which borrows from both the "large sample theory"

tradition in mathematical statistics and the "factor analysis" tradition in psychometrics, is that of *essential independence*[2].

For any (infinite) sequence of dichotomous items $\underline{X} = (X_1, X_2, X_3, \ldots)$, define *bounded item scores* to be functions $A_j(X_j)$ such that for some $M < \infty$, $|A_j(X_j)| \leq M$ for all j. We will call a bounded item score an *ordered item score* if moreover $A_j(0) \leq A_j(1)$. Define a *bounded test score* to be the average of the first J bounded item scores $\overline{A}_J = \frac{1}{J}\sum_{j=1}^{J} A_j(X_j)$. Finally, we will say the ordered item scores are *asymptotically discriminating* if $\frac{1}{J}\sum_{j=1}^{J}\{A_j(1) - A_j(0)\}$ is positive and bounded away from 0 as $J \to \infty$.

The infinite item sequence $\underline{X}$ is *essentially independent* (EI) *with respect to* $\underline{\Theta}$ if and only if

$$\lim_{J\to\infty} \mathrm{Var}\,(\overline{A}_J | \underline{\Theta} = \underline{\theta}) = 0 \tag{EI}$$

for all bounded test scores $\overline{A}_J$. When EI holds, $\overline{A}_J$ is a consistent estimator of the "true score" $\overline{A}_J(\underline{\theta}) = E[\overline{A}_J | \underline{\Theta} = \underline{\theta}]$, as $J \to \infty$. In particular, for a sequence of dichotomous items $\underline{X}$, EI implies that the proportion correct score $\overline{X}_J$ consistently estimates values of the *test characteristic function* $\overline{P}_J(\underline{\theta})$, as $J \to \infty$.

The item sequence $\underline{X}$ is *essentially unidimensional,* for which we shall write $d_E = 1$, if and only if (a) $\underline{X}$ is EI with respect to a unidimensional Θ; and (b) the items are *locally asymptotically discriminating*, LAD: for every set of ordered, asymptotically discriminating item scores, the "true score" $\overline{A}_J(\theta)$ is nondecreasing in θ, in the strong sense that to every θ there corresponds an interval N_θ containing θ and an $\epsilon_\theta > 0$ such that

$$\frac{\overline{A}_J(t) - \overline{A}_J(\theta)}{t - \theta} \geq \epsilon_\theta,\ \forall\, t \in N_\theta,\ t \neq \theta,\ \forall\, J. \tag{LAD}$$

If no such unidimensional Θ exists, we write $d_E > 1$.

It is not apparent from the above discussion, but Stout (1987) and Junker (1988, Section 3.2), make it clear that $d_E = 1$ can be checked from the marginal distribution of $\underline{X}$. A statistical procedure for testing the hypothesis that a set of J items $\underline{X}_J$ comes from a $d_E = 1$ item sequence has been developed by Stout (1987) and refined by Stout and Nandakumar (Nandakumar, 1987, 1990; Nandakumar and Stout, 1990).

When $d_E = 1$, $\overline{A}_J(\theta)$ may be inverted to produce estimates of θ directly:

[2] Actually, we use Stout's *strong essential independence*, with some minor changes in terminology to match Junker (1991). Any of the three variations of EI could be used; cf. Corollary 2.1 in Section 2.3.

Theorem 2.2 *(Stout, 1990). If the item sequence $\underline{X}$ satisfies $d_E = 1$ with respect to Θ then for any set of asymptotically discriminating item scores,*

$$\forall\, \epsilon > 0, \lim_{J\to\infty} P\left[|\overline{A}_J^{-1}(\overline{A}_J) - \theta| \le \epsilon \middle| \Theta = \theta\right] = 1, \tag{4}$$

where $\overline{A}_J^{-1}(u)$ is the inverse function for the "true score" $\overline{A}_J(\theta)$.

Indeed, under the conditions of Theorem 2.2 and some mild smoothness conditions, the maximum likelihood estimate of θ calculated as though LI were true is also consistent for θ (Junker, 1991). Moreover the latent trait with respect to which $d_E = 1$ holds is unique, up to a monotone transformation (Stout, 1990). It is valuable to think of EI as the greatest possible weakening of LI under which LI-based trait estimation/prediction schemes could be expected to work. In this sense, the study of EI is the study of robustness of ability estimators to variations from an LI latent structure. Clarke and Junker (in progress) pursue this matter in a more general setting.

2.3 Combining CA and $d_E = 1$

The following lemma tells us that under $d_E = 1$, for any finite subtest $\underline{Y}$ in $\underline{X}_J$, we may approximate expected values of the form $E[f(\underline{Y})|\Theta]$ with expected values of the form $E[f(\underline{Y})|\alpha_J \le \overline{X}_J \le \beta_J]$, as $J \to \infty$. We assume for the remainder of the paper that

$$P[C|\Theta = \theta] \text{ is continuous in } \theta \tag{5}$$

whenever C is an event depending on only finitely many X's. Condition (5) implies that $E[f(\underline{Y})|\Theta = t]$ is continuous in t for any function $f(\underline{Y})$ of finitely many items.

Lemma 2.1 *Suppose $\underline{X}$ satisfies* EI *and* LAD *with respect to some unidimensional Θ, and assume* (5). *If $f(\underline{Y})$ is a function which depends on only finitely many (fixed) items $\underline{Y} = (Y_1, \ldots, Y_{J_0})$ from $\underline{X}$, then for every set of bounded asymptotically discriminating item scores $A_j(X_j)$ and for each θ there exist $\epsilon_J \to 0$ for which*

$$\lim_{J\to\infty} E\left[f(\underline{Y}) \middle| |\overline{A}_J^{-1}(\overline{A}_J) - \theta| < \epsilon_J\right] = E[f(\underline{Y})\,|\Theta = \theta].$$

proof. We give the proof in the case that Θ is continuous, but a similar argument may be given for discrete numerical Θ.

For any event C, let 1_C take the value 1 if C is true and 0 if C is false, and let $E[f(\underline{Y}); C] = E[f(\underline{Y})1_C]$. We may decompose the expectation on the left above as

$$\begin{aligned} & E\left[f(\underline{Y})\,\middle|\,|\overline{A}_J^{-1}(\overline{A}_J) - \theta| < \epsilon\right] \\ = & \frac{E[f(\underline{Y}); |\Theta - \theta| < \epsilon]}{P[|\Theta - \theta| < \epsilon]} \cdot \frac{E[f(\underline{Y}); |\overline{A}_J^{-1}(\overline{A}_J) - \theta| < \epsilon]}{E[f(\underline{Y}); |\Theta - \theta| < \epsilon]} \cdot \frac{P[|\Theta - \theta| < \epsilon]}{P[|\overline{A}_J^{-1}(\overline{A}_J) - \theta| < \epsilon]} \\ = & \mathrm{I}(\epsilon) \cdot \mathrm{II}(\epsilon) \cdot \mathrm{III}(\epsilon). \end{aligned}$$

Note that for any rate $\epsilon = \epsilon_J \to 0$, $\mathrm{I}(\epsilon_J) \to E[f(\underline{Y})|\Theta = \theta]$, as $J \to \infty$, using the continuity condition (5) and the integral mean value theorem. The idea now is to choose $\epsilon = \epsilon_J \to 0$ so that $\mathrm{II} \to 1$ and $\mathrm{III} \to 1$ as $J \to \infty$. We will look at II explicitly; note that III is a special case of II. We have

$$\mathrm{II}(\epsilon) = 1 - \frac{E\{f(\underline{Y})[1_{\{|\overline{A}_J^{-1}(\overline{A}_J) - \theta| < \epsilon\}} - 1_{\{|\Theta - \theta| < \epsilon\}}]\}}{E\{f(\underline{Y})1_{\{|\Theta - \theta| < \epsilon\}}\}};$$

one can apply Theorem 2.2 to show that the numerator on the right tends to zero for each fixed $\epsilon > 0$ as $J \to \infty$; a simple diagonalization argument now yields a rate $\epsilon_J \to 0$ for which $\mathrm{II} \to 1$. A similar argument works for III, and a further diagonalization completes the proof. □

We can use Lemma 2.1 to gain information about the latent structure of an item sequence $\underline{X}$ from the manifest condition CA. Proposition 2.1 shows that CA and $d_E = 1$ together give the same local association condition (3) as $d_L = 1$ alone.

Proposition 2.1 *Suppose the item sequence* $\underline{X}$ *satisfies* CA *and* $d_E = 1$, *and suppose that* (5) *holds. Then (3) holds:*

$$\mathrm{Cov}\,(f(\underline{Y}), g(\underline{Y})|\Theta = \theta) \geq 0$$

for all θ, *all coordinatewise nondecreasing* f *and* g, *and all finite tests* $\underline{Y}$ *taken from* $\underline{X}$.

Remarks. By modifying the proof of Lemma 2.1 we could also conclude conditional association given $\Theta = \theta$, i.e., if $\underline{Z}$ were a finite test from $\underline{X}$ disjoint from $\underline{Y}$ and $h(\underline{Z})$ were any function, then

$$\mathrm{Cov}\,(f(\underline{Y}), g(\underline{Y})|h(\underline{Z}), \Theta = \theta) \geq 0.$$

proof. Let $\underline{Y} \subset \underline{X}_{J_0}$ be an arbitrary finite test, for fixed J_0, let $\underline{W} = (X_{J_0+1}, X_{J_0+2}, \ldots)$ and, for this proof, let $\overline{P}_J(\theta) = E[\frac{1}{J}\sum_1^J W_j|\theta]$. Using CA and a sequence ϵ_J obtained from Lemma 2.1,

$$\begin{aligned} 0 &\leq \operatorname{Cov}\left[f(\underline{Y}), g(\underline{Y}) \middle| |\overline{P}_J^{-1}(\overline{W}_J) - \theta| < \epsilon_J\right] \\ &\to \operatorname{Cov}\left(f(\underline{Y}), g(\underline{Y})|\Theta = \theta\right) \end{aligned}$$

as $J \to \infty$. □

Let us digress briefly to indicate another way in which CA and $d_E = 1$ interact well. Two alternative definitions of EI have been proposed by Stout (1990), one involving the full sequence $\underline{X}$ but taking absolute values of covariances,

$$\lim_{J\to\infty} \binom{J}{2}^{-1} \sum\sum_{1\leq i<j\leq J} |\operatorname{Cov}(X_i, X_j|\theta)| = 0, \tag{6}$$

and another involving "nonsparse subtests" which, in the present context, is equivalent to considering only those asymptotically discriminating item scores for which $A_J(0) \equiv 0$ and $A_J(1) \in \{0, 1\}$, and requiring

$$\lim_{J\to\infty} \binom{J}{2}^{-1} \sum\sum_{1\leq i<j\leq J} \operatorname{Cov}(A(X_i), A(X_j)|\theta) = 0. \tag{7}$$

It is not known in general whether these three definitions are equivalent. However, under CA they are:

Corollary 2.1 *If* CA *and* LAD *hold for* $\underline{X}$, *then all three definitions of* EI *are equivalent.*

proof. Condition (6) implies EI as defined in Section 2, which in turn implies (7), since each condition is a special case of the preceding one. For the converse directions, observe that if LAD holds (for the restricted case of $A_j(X_j) \in \{0, 1\}\ \forall j$), then by Proposition 2.1, $\operatorname{Cov}(X_i, X_j|\theta) \geq 0, \forall i, j$. In this case, (6) becomes a special case of (7), and we are done. □

Returning to our main development, the next proposition complements Proposition 2.1 by characterizing LI in terms of quantities that could in principle be approximated by manifest quantities $E[f(\underline{Y})|\alpha_J \leq \overline{X}_J \leq \beta_J]$ as in Lemma 2.1 under EI and LAD.

Proposition 2.2 *The item sequence* $\underline{X}$ *satisfies* LI *with respect to* Θ *if and only if the following two conditions hold:*

For all θ, all nondecreasing f and g, and all finite tests $\underline{Y}$ taken from $\underline{X}$,

$$\mathrm{Cov}\,(f(\underline{Y}), g(\underline{Y})|\Theta = \theta) \geq 0; \tag{8}$$

For all θ, i, and j,

$$\mathrm{Cov}\,(X_i, X_j|\Theta = \theta) \leq 0. \tag{9}$$

Remarks. Note that (8) says that $\underline{X}$ is *associated* (as defined by Esary, Proschan and Walkup, 1967), given $\Theta = \theta$, for all θ—indeed (8) is exactly the same as (3).

proof. That LI implies (8) follows from Esary, Proschan and Walkup (1967); (9) is trivially satisfied under LI with $\mathrm{Cov}\,(X_i, X_j|\theta) \equiv 0$. For a proof of the converse, in unconditional form, see Newman and Wright (1981) or Joag-Dev (1983). □

3 Two new conditions

Propositions 2.1 and 2.2 suggest that if we require both CA and $d_E = 1$ in the item sequence $\underline{X}$, we are not far away from a strictly unidimensional representation. Indeed, by Proposition 2.1, we know that CA and $d_E = 1$ will guarantee (8). In this section we will explore conditions on $P[\underline{X}_J = \underline{x}_J]$ that will guarantee (9) also. We also show that it is possible to check for monotone ICC's—a condition not explicitly provided for by CA and $d_E = 1$—by examining the "empirical ICC's" $P[X_j = 1|\overline{X}_J - X_j/J]$.

3.1 CA, $d_E = 1$, and useful unidimensional models

It can be seen from Theorem 2.1 and the discussion following the definition of essential independence that CA and EI are both necessary conditions for the sequence of items $\underline{X}$ to have a $d_L = 1$ representation. In this section, we consider several "thought examples" which suggest to what extent these two conditions suffice to characterize $d_L = 1$. We restrict our attention to *useful* $d_L = 1$ models, as defined in Section 1. Because LAD formalizes principle **U2** and Theorem 2.2 satisfies principle **U1**, any $d_E = 1$ model in which Θ varies in the examinee population is *useful* in the sense of Definition 1.1.

Example 3.1 CA *is not a guarantee that a useful* $d_L = 1$ *model exists.* Suppose $(X_1, X_2, \ldots)$ are independent (unconditionally). Then CA follows from a result of Esary, Proschan and Walkup (1967). Now, any Θ that we can estimate using principle **U1** will, according to the 0-1 Law of probability theory (e.g. Ash, 1972, p. 278), fail to vary from examinee to examinee. This violates principle **U3**; hence no *useful* latent trait model for this CA sequence exists. (Of course this is no surprise, since a fully independent sequence of response variables—e.g. coin flips—should intuitively not be able to tell us anything useful about a latent trait anyway!) □

Example 3.2 $d_E = 1$ *is not a guarantee that a useful* $d_L = 1$ *model exists.* Stout (1990, Example 2.3) gives a model for a sequence of "paragraph comprehension" questions which is a *useful* $d_E = 1$ model. We shall show that no *useful* $d_L = 1$ model can be formulated for this sequence of paragraph comprehension items. Indeed, if items X_i and X_j refer to the same reading passage we expect $\mathrm{Cov}\,(X_i, X_j | \Theta = \theta_o) \neq 0$ for some θ_o. Now suppose, by way of contradition, that there exists a unidimensional latent trait τ with respect to which LI and LAD hold; in particular $\mathrm{Cov}\,(X_i, X_j | \tau) = 0$. Now by Stout's unique trait theorem (Theorem 3.3 of Stout, 1990), τ would be a monotone—indeed invertible—transformation of $\Theta, \tau = g(\Theta)$. But, taking $t_o = g(\theta_o)$,

$$\begin{aligned} 0 &= \mathrm{Cov}\,(X_i, X_j | \tau = t_o) \\ &= \mathrm{Cov}\,(X_i, X_j | g(\Theta) = g(\theta_o)) \\ &= \mathrm{Cov}\,(X_i, X_j | \Theta = \theta_o) \\ &\neq 0. \end{aligned}$$

This contradiction shows that no such τ can exist, i.e. no *useful* $d_L = 1$ model exists for the sequence of paragraph comprehension items. □

Example 3.3 CA *and* $d_E = 1$ *together may not guarantee that a useful* $d_L = 1$ *model exists.* Let $\underline{X}$ be an item sequence satisfying $d_E = 1$ with respect to s me unidimensional Θ. We may imagine the Θ in this $d_E = 1$ representation as being the first coordinate of the latent trait vector $\underline{\Theta} = (\Theta, \Theta_2, \Theta_3, \ldots, \Theta_d)$ needed for a $d_L = d$ representation:

$$P[\underline{X}_J = \underline{x}_J | \Theta = \theta] = \int P[\underline{X}_J = \underline{x}_J | \underline{\Theta} = \underline{\theta}] f(\underline{\theta} | \Theta = \theta) d\underline{\theta}$$

for each J, where, by LI with respect to $\underline{\Theta} = (\Theta, \Theta_2, \Theta_3, \ldots, \Theta_d)$,

$$P[\underline{X}_J = \underline{x}_J | \underline{\Theta} = \underline{\theta}] = \prod_{j=1}^{J} P_j(\underline{\theta})^{x_j}(1 - P_j(\underline{\theta}))^{1-x_j}.$$

In general, although $d_E = 1$ means that

$$\lim_{J\to\infty} \binom{J}{2}^{-1} \sum\sum_{1\le i<j\le J} \mathrm{Cov}\,(Z_i, Z_j | \Theta = \theta) = 0$$

for every nonsparse subsequence $\underline{Z}$ from $\underline{X}$, the individual covariances

$$\mathrm{Cov}\,(Z_i, Z_j | \Theta = \theta) = \mathrm{Cov}\,(P_i(\underline{\Theta}), P_j(\underline{\Theta}) | \Theta = \theta)$$

may be positive or negative, depending on how the traits $\Theta_2, \Theta_3, \ldots, \Theta_d$ interact with Θ.

Now suppose $\underline{X}$ satisfies CA also. Then by Proposition 2.1,

$$\mathrm{Cov}\,(X_i, X_j | \Theta = \theta) \ge 0, \forall\, i \ne j; \tag{10}$$

in fact the stronger condition (3) holds. Thus, not only is Θ the dominant trait for $\underline{X}$, but the "minor traits" needed for LI to hold are concordant with Θ, in the sense that they interact with Θ so as to keep the local inter-item covariances nonnegative.

Under CA and $d_E = 1$, therefore, there is enough coherence among the items that covariances between items, given $\Theta = \theta$, are nonnegative. Indeed it is quite plausible that under these conditions, for some tests, some of the inequalities in (10) will be strict (despite its plausibility we have not been able to construct an example in which this may rigorously be shown). But if any of the inequalties (10) are strict for a dominant latent trait Θ with respect to which $d_E = 1$, then there cannot exist any other unidimensional trait τ with respect to which a useful $d_L = 1$ model exists; this follows by the same argument as in Example 3.2. □

3.2 CSN: A natural negative covariance condition

Thus some condition in addition to CA and $d_E = 1$ seems to be needed to get a useful $d_L = 1$ model. When LI holds, we know that (3) holds also: $\mathrm{Cov}\,(f(\underline{Y}), g(\underline{Y}) | \Theta = \theta) \ge 0$ for all finite subtests $\underline{Y}$ and nondecreasing functions f and g. As indicated in Example 3.3, there

may be some situations in which EI and CA hold but the implied coherence among items is so tight that LI cannot also hold. What is needed is a condition that loosens this coherence so that, despite (3), individual item pairs have zero covariance.

An interesting condition derived by Joag-Dev and Proschan (1982) implies that when LI with respect to Θ does hold, then

$$\mathrm{Cov}\,(X_i, X_j | \overline{X}_J, \Theta) \leq 0 \tag{LCSN}$$

for all $i < j \leq J$. This says that, under LI, the items are "not too tightly bound together" even though (2) holds: each X_i and X_j are sufficiently free of one another among examinees of the same ability that when X_i increases from one examinee to the next, X_j is free to decrease so that the test score $\overline{X}_J$ may be kept constant. The abbreviation LCSN stands for *locally, covariances given test scores are negative.*

However, LCSN is a condition on the latent, not the manifest, structure. To obtain a natural manifest structure analogue to LCSN it is useful to consider the special case of the locally independent Rasch model. Here $\overline{X}_J$ is sufficient for Θ, i.e. $(X_1, X_2, \ldots, X_J)$ are independent of Θ given $\overline{X}_J$. One consequence of this is that

$$\mathrm{Cov}\,(X_i, X_j | \overline{X}_J) = \mathrm{Cov}\,(X_i, X_j | \overline{X}_J, \Theta), \tag{11}$$

in which case LCSN is equivalent to the manifest condition

$$\mathrm{Cov}\,(X_i, X_j | \overline{X}_J) \leq 0 \tag{CSN}$$

for all $i < j \leq J$. Here CSN should be read as *covariances given test scores are negative.* In practice it may be necessary to allow (CSN) to be violated for very small values of J, e.g. $J < 10$.

Because $\overline{X}_J$ is not sufficient for Θ outside the Rasch model, CSN is an imperfect substitute for LCSN. It is valuable to know how closely related LCSN and CSN are under LI. Suppose LI and hence LCSN holds. To examine CSN one might consider the decomposition

$$\mathrm{Cov}\,(X_i, X_j | \overline{X}_J) = E\left[\mathrm{Cov}\,(X_i, X_j | \overline{X}_J, \Theta) \middle| \overline{X}_J\right] + \mathrm{Cov}\,\left[E[X_i | \overline{X}_J, \Theta], E[X_j | \overline{X}_J, \Theta] \middle| \overline{X}_J\right].$$

The first term on the right is nonpositive, by LCSN. The second term may be negative or positive, but should be small since under $d_L = 1$ the (posterior) distribution of Θ given $\overline{X}_J$

should have very low variance, as J grows. Some preliminary work of B. Clarke and J. K. Ghosh (Clarke, priv. comm.) points toward a proof of this assertion. This suggests that in most $d_L = 1$ models

$$\mathrm{Cov}\,(X_i, X_j|\overline{X}_J) \approx E\left[\mathrm{Cov}\,(X_i, X_j|\overline{X}_J, \Theta)\middle|\overline{X}_J\right] \tag{12}$$

for longer tests, even though $\overline{X}_J$ is not sufficient for Θ. In particular when $\mathrm{Cov}\,(X_i, X_j|\overline{X}_J)$ fails to be negative for a LI model, we at least expect it to be near zero. Thus one might consider, not CSN, but rather a condition like

$$\mathrm{Cov}\,(X_i, X_j|\overline{X}_J) \leq 0 + \epsilon_J \tag{13}$$

for suitably chosen $\epsilon_J > 0$; this is similar to including an indifference region in a test of the null hypothesis CSN against a general alternative. The plausibility of CSN and (13) in locally independent two parameter and three parameter logistic IRT models is illustrated in a small simulation in Appendix C.

3.3 MM: Manifest monotonicity

When one desires to check M in practical situations, a condition like the following is often used. Let $\overline{X}_{iJ} = \overline{X}_J - X_i/J$; we will say *manifest monotonicity*, MM, holds if

$$E[X_i|\overline{X}_{iJ}] \textit{ is nondecreasing in } \overline{X}_{iJ} \tag{MM}$$

for all $i \leq J$ (and all J). This intuitively appealing monotonicity check is intimately related to $d_L = 1$ latent structure, in that

(a) LI, M $\Rightarrow$ MM;

(b) EI, LAD, MM $\Rightarrow$ M; and hence

(c) under LI and LAD, MM $\Leftrightarrow$ M.

Assertions (a) and (b) are proved in Proposition 3.1, and (c) is an immediate corollary. Molenaar (priv. comm.) has independently discovered (a), and we report examples of Molenaar and Snijder in Appendix B that indicate the limitations of the method of proof for (a).

Lemma 3.1 *If* LI *and* M *hold for an item sequence* $\underline{X}$ *and latent trait* Θ*, then* Θ *is stochastically increasing in* $S = \overline{X}_{iJ}$*:*

$$\forall\ a \leq b\ \forall\ c : P[\theta > c|S = a] \leq P[\theta > c|S = b] \tag{14}$$

whenever the conditional probabilities are defined.

proof. We may apply Grayson (1988), Theorem 2, to the subtest $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_J)$ to see that the score $S = \overline{X}_{iJ}$ has the monotone likelihood ratio property

$$R_{ab}(\theta) \equiv \frac{P[S = b|\theta]}{P[S = a|\theta]} \textit{ nondecreasing in } \theta,\ \forall\ a \leq b, \tag{15}$$

whenever the conditional probabilities are defined. To establish (14) (for any score S satisfying (15)), we may write its left hand side as

$$\begin{aligned} P[\theta > c|S = a] &= \frac{\int_c^{+\infty} P[S = a|\theta]dF(\theta)}{\int_{-\infty}^{+\infty} P[S = a|\theta]dF(\theta)} \\ &= P[T > c] \end{aligned}$$

where T is a random variable with density proportional to $P[S = a|t] \cdot dF(t)$. (I.e., $T = [\Theta|S = a]$.) On the other hand, the right hand side of (14) may be written as

$$\begin{aligned} P[\theta > c|S = b] &= \frac{\int_c^{+\infty} P[S = b|\theta]dF(\theta)}{\int_{-\infty}^{+\infty} P[S = b|\theta]dF(\theta)} \\ &= \frac{\int_c^{+\infty} P[S = a|\theta]R_{ab}(\theta)dF(\theta)}{\int_{-\infty}^{+\infty} P[S = a|\theta]R_{ab}(\theta)dF(\theta)} \\ &= \frac{E[R_{ab}(T)1_{\{T>c\}}]}{E[R_{ab}(T)]} \end{aligned}$$

for the same random variable T (where 1_C is the function that takes the value 1 when C is true and 0 otherwise). Hence (14) is equivalent to the assertion that

$$P[T > c] \cdot E[R_{ab}(T)] \leq E[R_{ab}(T)1_{\{T>c\}}]$$

which follows from property (P3) of Esary, Proschan and Walkup (1967), since $g(T) = 1_{\{T>c\}}$ and $h(T) = R_{ab}(T)$ are both nondecreasing functions of T. □

Proposition 3.1

(a) LI, M $\Rightarrow$ MM

(b) EI, LAD, MM *and* (5) $\Rightarrow$ M

proof. (a) Note that

$$\begin{aligned} E[\overline{X}_i|\overline{X}_{iJ}] &= E\left[E[\overline{X}_i|\overline{X}_{iJ},\Theta]\Big|\overline{X}_{iJ}\right] \\ &= E[P_i(\Theta)|\overline{X}_{iJ}] \end{aligned} \tag{16}$$

by LI. This last expectation is nondecreasing in $\overline{X}_{iJ}$ by M and Lemma 3.1, using a result of Lehmann (1955).

(b) Let $\theta^{(1)} < \theta^{(2)}$; then there exist sequences $\alpha_J^{(1)} \le \beta_J^{(1)}$ and $\alpha_J^{(2)} \le \beta_J^{(2)}$ with $\beta_J^{(1)} < \alpha_J^{(2)}$ for all large J, such that $\{\alpha_J^{(i)} \le \overline{X}_{iJ} \le \beta_J^{(i)}\} = \{|\overline{P}_J^{-1}(\overline{X}_{iJ}) - \theta^{(i)}| < \epsilon_J^{(i)}\}$ from Lemma 2.1. Then

$$\begin{aligned} E[X_i|\Theta = \theta^{(1)}] &= \lim_{J\to\infty} E[X_i|\alpha_J^{(1)} \le \overline{X}_{iJ} \le \beta_J^{(1)}] \\ &\le \lim_{J\to\infty} E[X_i|\alpha_J^{(2)} \le \overline{X}_{iJ} \le \beta_J^{(2)}] \\ &= E[X_i|\Theta = \theta^{(2)}] \end{aligned}$$

where the middle inequality follows from the fact that

$$\begin{aligned} E[X_i|\alpha_J^{(1)} \le \overline{X}_{iJ} \le \beta_J^{(1)}] &= \frac{P[X_{i=1}, \alpha_J^{(1)} \le \overline{X}_{iJ} \le \beta_J^{(1)}]}{P[\alpha_J^{(1)} \le \overline{X}_{iJ} \le \beta_J^{(1)}]} \\ &= \frac{\sum_{c=\alpha_J^{(1)}}^{\beta_J^{(1)}} E[X_i|\overline{X}_{iJ}=c]P[\overline{X}_{iJ}=c]}{\sum_{c=\alpha_J^{(1)}}^{\beta_J^{(1)}} P[\overline{X}_{iJ}=c]} \\ &\le \frac{\sum_{c=\alpha_J^{(2)}}^{\beta_J^{(2)}} E[X_i|\overline{X}_{iJ}=c]P[\overline{X}_{iJ}=c]}{\sum_{c=\alpha_J^{(2)}}^{\beta_J^{(2)}} P[\overline{X}_{iJ}=c]} \\ &= E[X_i|\alpha_J^{(2)} \le \overline{X}_{iJ} \le \beta_J^{(2)}] \end{aligned}$$

(under MM, the second ratio of sums above is a weighted average of larger conditional probabilities than the first one). □

Corollary 3.1 *Under* LI, LAD *and* (5) *we have*

$$\text{MM} \Leftrightarrow \text{M}$$

4 Characterization of $d_L = 1$

The previous two sections may be summarized in the following theorem.

Theorem 4.1 *If $\underline{X}$ is an item seqence satisfying $d_L = 1$ and if* LAD *holds with respect to the latent trait Θ, then each of the conditions* CA, $d_E = 1$, LCSN *and* MM *hold.*

We show in this section that the converse is also true: the four conditions CA, $d_E = 1$, LCSN and MM guarantee a useful $d_L = 1$ representation; hence these conditions characterize useful $d_L = 1$ representations. Moreover the converse implication is still true if LCSN is replaced with its manifest analogue condition CSN.

The main problem with obtaining these two converses of Theorem 4.1 is that our negative covariance criteria CSN and LCSN use conditioning on fixed values of $\overline{X}_J$ only, whereas Lemma 2.1 gives approximations to (8) and (9) which require conditioning on intervals $\alpha_J \leq \overline{X}_J \leq \beta_J$. The next lemma connects these two forms of conditioning. We will assume that for each J and $i \leq J$ there exist differentiable g_{iJ} such that

$$\left.\begin{array}{l} E[X_i|\overline{X}_J] \equiv g_{iJ}(\overline{X}_J) \\ \sup_{i,J,u} |g'_{iJ}(u)| \leq M < \infty \end{array}\right\} \tag{17}$$

and that for each $J, i \leq J$, and θ there exist differentiable $g_{iJ\theta}$ such that

$$\left.\begin{array}{l} E[X_i|\overline{X}_J, \Theta = \theta] \equiv g_{iJ\theta}(\overline{X}_J) \\ \sup_{i,J,u} |g'_{iJ\theta}(u)| \leq M_\theta < \infty \end{array}\right\} \tag{18}$$

The conditions (17) and (18) are maximum discrimination conditions on item-test regressions; most likely they would be acceptable in practice. In particular, note that (17) and (18) do *not* hide a monotonicity assumption.

Lemma 4.1 *(a) Suppose* CSN *holds, and suppose* (17) *also holds. Then for any constants $\alpha_J \leq \beta_J$ for which the covariances are defined, and for which $\beta_J - \alpha_J \to 0$,*

$$\limsup_{j\to\infty} \operatorname{Cov}(X_i, X_j|\alpha_J \leq \overline{X}_J \leq \beta_J) \leq 0. \tag{19}$$

(b) Suppose LCSN *holds, and suppose* (18) *also holds. Then for any constants $\alpha_J \leq \beta_J$ for which the covariances are defined, and for which $\beta_J - \alpha_J \to 0$,*

$$\limsup_{J\to\infty} \mathrm{Cov}\,(X_i, X_j|\alpha_J \leq \overline{X}_J \leq \beta_J, \theta) \leq 0. \tag{20}$$

proof. We will do part (a) only; part (b) is virtually identical. We have

$$\begin{aligned} \mathrm{Cov}\,(X_i, X_j|\alpha_J \leq \overline{X}_J \leq \beta_J) &= E\left[\mathrm{Cov}\,(X_i, X_j|\overline{X}_J)\middle|\,\alpha_J \leq \overline{X}_J \leq \beta_J\right] \\ &\quad + \mathrm{Cov}\,\left[E[X_i|\overline{X}_J], E[X_j|\overline{X}_J]\middle|\,\alpha_J \leq \overline{X}_J \leq \beta_J\right]. \end{aligned} \tag{21}$$

The first term on the right is evidently nonpositive for J large. Let us drop the conditioning on $\alpha_J \leq \overline{X}_J \leq \beta_J$ from the notation for brevity; then the second term in (21) is

$$\begin{aligned} \mathrm{Cov}\,[E[X_i|\overline{X}_J], E[X_j|\overline{X}_J]] &= \mathrm{Cov}\,[g_{iJ}(\overline{X}_J), g_{iJ}(\overline{X}_J)] \\ &\leq \{\mathrm{Var}\, g_{iJ}(\overline{X}_J) \cdot \mathrm{Var}\, g_{jJ}(\overline{X}_J)\}^{\frac{1}{2}}, \end{aligned}$$

by the Cauchy-Schwarz inequality. Now applying Taylor's theorem,

$$\mathrm{Var}\, g_{iJ}(\overline{X}_J) \quad \leq \quad [\max_u |g'_{iJ}(u)|]^2 \cdot \mathrm{Var}\, \overline{X}_J$$

so that conditioning on $\alpha_J \leq \overline{X}_J \leq \beta_J$, which forces $\mathrm{Var}\, \overline{X}_J \to 0$ and hence $\mathrm{Var}\, g_{iJ}(\overline{X}_J) \to 0$ as $J \to \infty$, also forces the second term in (21) to go to zero, completing the proof. □

Now we are ready to state and prove the two converses to Theorem 4.1. The formal statement of our characterization of useful $d_L = 1$ models is

Theorem 4.2 *Suppose $\underline{X}$ is an item sequence and Θ is a unidimensional trait, and suppose* (5), (17) *and* (18) hold. Then

(a) CA, $d_E = 1$, LCSN, MM $\Leftrightarrow$ $d_L = 1$, LAD

(b) CA, $d_E = 1$, CSN, MM $\Rightarrow$ $d_L = 1$, LAD

Remarks. In the implications "$\Rightarrow$" in (a) and (b), Θ is the trait with respect to which $d_E = 1$ holds, and the theorem asserts that in fact $d_L = 1$ holds with respect to this Θ. In the implication "$\Leftarrow$" in (a), Θ is the trait with respect to which $d_L = 1$ holds. In both cases, Θ is unique up to monotone transformation (Stout, 1990, Theorem 3.3).

proof. It is more convenient to prove (b) first.

Part (b), "$\Rightarrow$": There are three conditions to check on the right: LI, M, and LAD. LAD follows from $d_E = 1$ by definition. M follows from MM and $d_E = 1$ via Proposition 3.1(b). LI follows from CA, $d_E = 1$ and LCSN, using Proposition 2.1, Proposition 2.2, and Lemma 4.1(a), since (19) implies that under CSN and $d_E = 1$ we have $\mathrm{Cov}\,(X_i, X_j|\theta) \leq 0$ for all i, j and θ.

Part (a), "$\Rightarrow$": Again we must check LI, M and LAD. LAD and M follow as before. LI follows again, using Proposition 2.1, Proposition 2.2, and Lemma 4.1(b), since now (20) implies that under LCSN $\mathrm{Cov}\,(X_i, X_j|\theta) \leq 0$ for all i, j, and θ (a conditional [given $\Theta = \theta$] form of Lemma 2.1 is needed to show this, but this is straightforward).

Part (a), "$\Leftarrow$": This is Theorem 4.1, but we state the proof for completeness. We must check MM, CA, EI, LAD and LCSN. MM and CA follow from $d_L = 1$ by Proposition 3.1(a) and Theorem 2.1, respectively. EI follows from LI trivially, LAD is assumed on the right, and LCSN follows from LI via a conditional form of Theorem 2.8 of Joag-Dev and Proschan (1982). □

Theorem 4.2(a) gives a complete characterization of $d_L = 1$ representations among "smooth" representations satisfying the mild monotonicity condition LAD—this is essentially the class of *useful* $d_L = 1$ representations, assuming that the distribution of Θ is not concentrated at one point. The characterization is of interest because the conditions MM, CA and EI, are all conditions on the manifest structure $P[\underline{X}_J = \underline{x}_J]$ of $\underline{X}$. LCSN is not itself a "manifest" condition, but it appears to be quite close to its natural manifest structure analogue CSN in practice. Theorem 4.2(b) gives reasonably general conditions on the manifest structure of $\underline{X}$ which are sufficient to guarantee a "useful" $d_L = 1$ representation. Note that overly restrictive assumptions, such as detailed knowledge of the forms of the ICC's or of the distribution of Θ, are not needed in this approach.

It is important to point out that Theorem 4.2 is possible only through the use of an infinite item sequence $\underline{X}$. The constraints put on the latent structure of a finite set of items $\underline{X}_J = (X_1, \ldots, X_J)$ by the distribution of $\underline{X}_J$ are not strong enough, in general, to guarantee a particular form for the latent structure. The principal difference between the finite case $\underline{X}_J$ and the infinite case $\underline{X}$ is that, whereas a latent trait may be only imperfectly estimated using $\underline{X}_J$, it can be known with complete accuracy using $\underline{X}$ (see Levine, 1985, for a related

discussion about the limits of our ability to know Θ from $\underline{X}_J$ for finite J). Thus the use of (conceptually) infinite sequences of items seems absolutely vital to clarify model-building and model-identification issues (see also Stout, 1990, for a discussion of this point).

5 Discussion

In this paper we have combined the conditional association (CA) approach of Holland and Rosenbaum and the essential unidimensionality ($d_E = 1$) approach of Stout to produce a nearly complete nonparametric characterization of *useful* $d_L = 1$ representations for dichotomous IRT data. The three principles **U1**, **U2** and **U3** for a *useful* representation require that the latent trait Θ can be consistently estimated from the item responses; that Θ is monotonically related to the test's "true score;" and that Θ is not constant in the examinee population.

In Section 2 we reviewed the CA and $d_E = 1$ conditions. A crucial feature of our analysis was the embedding of the finite-length test $\underline{X}_J$ into an infinite sequence of items $\underline{X}$ which extends the features of the observed set of items $\underline{X}_J$. This embedding is needed for two reasons: first, several authors have observed that estimation of traits to arbitrary precision—and hence identification of latent structure—cannot be expected to work unless the test length is allowed to grow without bound; and second, the embedding is needed to discuss Stout's notions of essential unidimensionality. Moreover in many practical settings, the embedding can be justified as a continuation of the usual process used to manufacture items (c.f. Stout, 1990). Under CA, all three definitions of $d_E = 1$ proposed by Stout are equivalent, and CA and $d_E = 1$ together bring us close to a $d_L = 1$ representation.

In Section 3 we developed two further conditions which seem to be needed to obtain a characterization of strict unidimensionality. First, the negative covariance condition CSN ($\mathrm{Cov}(X_i, X_j|\overline{X}_J) \leq 0$) is a natural one to add to CA and $d_E = 1$ to ensure that the items are locally independent with respect to Θ. CSN is guaranteed to be true in the Rasch model, and a simulation in Appendix C shows that CSN is also plausible in two-parameter and three-parameter logistic models. Second, monotonicity of the empirical ICC's $P[X_j = 1|\overline{X}_J - X_j/J]$ is intimately related to ICC monotonicity: this "manifest monotonicity" must hold if $d_L = 1$ holds; and conversely it can be used to verify monotonicity of the usual ICC's

$P_j(\theta)$ when $d_E = 1$ holds.

In Section 4 we showed that *useful* $d_L = 1$ representations may be characterized by the conditions CA, $d_E = 1$, manifest monotonicity, and a local version of CSN. If the local CSN condition is replaced with CSN itself, we obtain a fairly general set of sufficient conditions on the manifest structure $P[\underline{X}_J = \underline{X}_J]$ as $J \to \infty$, for a useful $d_L = 1$ representation to hold. These conditions have the agreeable property that parametric forms of the ICC's are not needed to check $d_L = 1$.

It is important to note that some form of monotonicity or ICC smoothness is needed to avoid meaningless models; an example in Appendix A, due to Suppes and Zanotti (1981), illustrates this. Our preferred "nonparametric" condition has been Stout's local asymptotic discrimination LAD condition. In parametric settings, a general monotonicity condition such as LAD might be dropped in the face of other smoothness available from the parametric form of the model. However such a condition is often plausible, even if the ICC's are not monotone, and greatly enhances the interpretability of the model.

The results of this paper bring us quite close to a characterization of useful $d_L = 1$ representations solely in terms of the distribution of the manifest item response data. This is valuable for two reasons. First, it is hoped that presenting them stimulates further discussion of the basic components of IRT modeling. For example, the local form of CSN in the characterization above suggests that both a positive covariance condition like CA, *and a negative covariance condition* like CSN seem needed to properly understand the general $d_L = 1$ assumptions. Second, such a characterization suggests practical nonparametric tests of fit for the general $d_L = 1$ representation. But the practical application of the characterization theorem requires one to deal with a potentially difficult multiple inference problem: testing CSN involves combining $J \cdot \binom{J}{2}$ statistical tests, and CA involves even more statistical tests. Understanding this multiple-inference problem and exploiting possible dependencies among the tests to reduce the problem is an important future step in this research.

Acknowledgements

A preliminary version of this material appeared as Chapter 4 of the author's Ph. D. dissertation. This work was greatly facilitated by enjoyable and stimulating discussions with Bill Stout. Comments on parts of this work by Bertrand Clarke and Paul Holland were very much appreciated. Finally, the author is grateful to Ivo Molenaar for allowing the examples in Appendix B to appear in this report.

A The "necessity" of LI, M and D

The assumptions LI, M and D can be—and sometimes are—*weakened* but not *dropped*: if any one of them is completely dropped, the resulting version of (1) can be made to fit any distribution of dichotomous items. Two of the examples which show this are known in the literature, but we repeat them here for completeness. Example A.3 appears to be new.

Example A.1 *$d = 1$ and LI hold, but M dropped.* (Suppes and Zanotti, 1981). Here we allow the ICC's to be arbitrarily rough and nonmonotone; our goal is to represent an arbitrary distribution $P[\underline{X}_J = \underline{x}_J]$ as in (1). Let $\Theta = \sum_1^J 2^{-j} X_j$, i.e. Θ is the base-two fraction $0.X_1X_2\ldots X_J$. Then the ICC's may be written as

$$P[X_j = 1|\theta] = 1_{\{\mathrm{Int}\,(2^j\cdot\theta) \bmod 2=1\}},$$

where "Int (t)" is the integer part of t, and 1_C equals 1 when C is true and 0 when C is false. The likelihood may be written as usual as under LI, and the distribution of θ is described by a probability mass function

$$f(\theta) = \begin{cases} P[\underline{X}_J = \underline{x}_J], & \text{if } \exists\, \underline{x}_J\text{: } \theta = \sum_1^J 2^{-j} x_j; \\ 0, & \text{if } \nexists\, \underline{x}_J. \end{cases}$$

Finally, (1) becomes

$$P[\underline{X}_J = \underline{x}_J] = \sum_{\theta: f(\theta)\neq 0} \prod_{j=1}^{J} P_j(\theta)^{x_j}(1 - P_j(\theta))^{1-x_j} f(\theta).$$

This example can be modified to allow infinitely many items, items with more than two response categories, and items with continuous responses. □

Example A.2 *LI and M hold, but $d = J$.* (Holland and Rosenbaum, 1986; Stout, 1990). To write the representation (1) in this setting, we put $\underline{\Theta} = \underline{X}_J$, so that each component $\theta_j \equiv x_j$ records exactly the response to the j^{th} item. The ICC's may be written as

$$P[X_j = 1|\underline{\theta}] = 1_{\{\theta_j=1\}},$$

which are certainly coordinatewise nondecreasing in $\underline{\theta}$; the multivariate distribution of $\underline{\Theta}$ is exactly the same as for $\underline{X}_J$:

$$f(\underline{\theta}) = P[\underline{\Theta} = \underline{\theta}] = P[\underline{X}_J = \underline{\theta}].$$

Here, (1) becomes

$$P[\underline{X}_J = \underline{x}_J] = \sum_{\underline{\theta}} \prod_{j=1}^{J} P_j(\underline{\theta})^{x_j}(1 - P_j(\underline{\theta}))^{1-x_j} f(\underline{\theta}). \qquad \square$$

Example A.3 *$d = 1$ and M hold, but LI dropped.* Here we set $\Theta = \prod_1^J X_j$. The ICC's

$$P[X_j = 1|\theta] = \begin{cases} 1, & \text{if } \theta = 1; \\ 0, & \text{if } \theta = 0 \end{cases}$$

remain monotone, but the joint likelihood cannot be written as under LI. Instead,

$$P[\underline{X}_J = \underline{x}_J|\theta] = \begin{cases} P[\underline{X}_J = \underline{x}_J], & \text{if } \theta = 0 \text{ and } \sum_1^J x_j < J; \\ P[\underline{X}_J = \underline{x}_J], & \text{if } \theta = 1 \text{ and } \sum_1^J x_j = J; \\ 0, & \text{otherwise} \end{cases}$$

and the probability mass function for θ is

$$f(\theta) = P[\textstyle\sum_1^J X_j = J]^{\theta} P[\sum_1^J X_j < J]^{(1-\theta)}.$$

Finally, the representation (1) is simply written as

$$P[\underline{X}_J = \underline{x}_J] = \sum_{\theta=0}^{1} P[\underline{X}_J = \underline{x}_J|\theta] f(\theta). \qquad \square$$

It is worth remarking that each of these examples is very much degenerate, as a piece of psychometric modeling. In Example A.1 and Example A.2 the latent variables chosen are

not latent at all, in that they are completely determined by the examinee's responses (and thus represent no broader sense of a psychological construct than the examinee's responses themselves). In Example A.3 the problem is the "opposite": here θ captures virtually nothing about the examinee's response behavior. The examples should be taken as indications of how quicky the modeling process can devolve if the assumptions LI, M and D are weakened too much.

B Limits on generalizing MM

Molenaar (priv. comm.) has independently discovered Proposition 3.1(a), and reports counterexamples indicating the limitations to extensions of this result. Example B.1 shows that the method of proof of Lemma 3.1 cannot be extended to the polytomous case. Example B.2 shows that, in Proposition 3.1(a), we cannot replace the "delete average" $\overline{X}_{iJ} = \frac{1}{J}\sum_1^J X_j - X_i/J$ with the more natural average over all items $\overline{X}_J$.

Example B.1 *(I. Molenaar). The monotone likelihood ratio property (15) of Grayson (1988) does not extend to polytomous items.* Let $0 \leq \Theta \leq 1$ and consider a single graded-response item X taking the three values 0, 1 or 2, with

$$P[X \geq 1|\theta] = \begin{cases} 3\theta, & 0 < \theta \leq 1/4 \\ \frac{3}{4} + \frac{1}{4}\theta, & 1/4 < \theta \leq 1 \end{cases}$$

$$P[X \geq 2|\theta] = \begin{cases} 0, & 0 \leq \theta \leq 1/4 \\ 3(\theta - \frac{1}{4}), & 1/4 < \theta \leq 1/2 \\ \frac{1}{2} + \frac{1}{2}\theta, & 1/2 < \theta \leq 1 \end{cases}$$

In this case, for $\theta_0 = 1/4$ and $\theta_1 = 1/2$, the likelihood ratio

$$\frac{P[X = x|\theta_0]}{P[X = x|\theta_1]} \tag{22}$$

is not monotone in x. For one may calculate

	$\theta_0 = 1/4$	$\theta_1 = 1/2$	Ratio
$P[X = 0\|\theta]$	1/4	1/8	2
$P[X = 1\|\theta]$	3/4	1/56	42
$P[X = 2\|\theta]$	0	6/7	0

Here the ratio (22) is not monotone in x. A consequence of (15) in the proof of Lemma 3.1 would be that the likelihood ratio (22) is monotone in x; since it is not, the proof of Lemma 3.1 cannot be extended to the polytomous case. □

Example B.2 *(T. Snijder). $P[X_j = 1|\overline{X}_J = s]$ need not increase with s, and hence we may not replace $\overline{X}_{iJ}$ with $\overline{X}_J$ in Proposition 3.1(a).* Consider three dichotomous item, and a two-point distribution for Θ, $P(\Theta = \theta_0) = P(\Theta = \theta_1) = \frac{1}{2}$. Let

$$\begin{aligned} P_j(\theta_0) &\equiv \epsilon,\ j = 1,2,3; \\ P_1(\theta_1) &= \frac{1}{2}; \\ P_2(\theta_1) &= 1 - \epsilon; \text{ and} \\ P_3(\theta_1) &= 1 - \epsilon. \end{aligned}$$

It follows that, as $\epsilon \to 0$,

$$\begin{aligned} P[X_1 = 1|\overline{X}_J = 1/3] &= \frac{(1-\epsilon)^2 + \frac{1}{2}\epsilon}{3(1-\epsilon)^2 + 1 - \frac{1}{2}\epsilon} \to \frac{1}{4} \\ P[X_1 = 1|\overline{X}_J = 2/3] &= \frac{\epsilon + 2\epsilon^2}{\frac{1}{2} + \frac{1}{2}\epsilon + 3\epsilon^2} \to 0. \end{aligned}$$

Hence for small $\epsilon > 0$, $P[X_1 = 1|\overline{X}_J = 1/3] > P[X_1 = 1|\overline{X}_J = 2/3]$. □

C An illustration of CSN

We have argued in Section 3.2 that the CSN condition

$$\mathrm{Cov}\,(X_i, X_j|\overline{X}_J) \leq 0$$

is a reasonable one to look for in dichotomous test data, to ensure that the local (given $\Theta = \theta$) association between items is not too strong, under CA, for LI to hold.

To illustrate the plausibility of CSN in common $d_L = 1$ IRT models, a small simulation study was performed. Item characteristic curves were taken to be of the logistic form

$$P_j(\theta) = c_j + (1 - c_j)\frac{1}{1 + \exp\{-1.7a_j(\theta - b_j)\}}. \tag{23}$$

All three of the popular ICC models based on (23) were considered: the Rasch model (one parameter logistic, 1PL) in which $c_j \equiv 0$ and $a_j \equiv 1$; the two parameter logistic model (2PL) in which $c_j \equiv 0$; and the three parameter logistic (3PL) model in which all parameters are free. Free parameters were generated as appropriate for each model as follows.

- $a_j = 0.5 + 0.5(j \bmod 3) + \alpha_j$, where α_j is $N(0, 0.0625)$ [normal, mean zero, variance 0.0625] noise, truncated so that $0.5 \leq a_j \leq 1.5$;
- $b_j = -2.0 + 4.0\frac{j-1}{J-1} + \beta_j$, where β_j is $N(0, \frac{1}{J^2})$ noise;
- $c_j = 0.2 + \gamma_j$, where γ_j is $N(0, 0.01)$ noise, truncated so that $0.0 \leq c_j \leq 0.4$.

For each of the three models, a 2000-examinee test administration was simulated, for varying test lengths J. Examinee abilities were sampled from a $N(0,1)$ distribution and responses were generated according to the corresponding locally independent IRT model.

The Mantel-Haenszel (M-H) one-sided z-test for

$$H_0\text{: } \mathrm{Cov}\,(X_i, X_j|\overline{X}_J) \leq 0 \text{ vs. } H_1\text{: } \mathrm{Cov}\,(X_i, X_j|\overline{X}_J) > 0$$

combined across $\overline{X}_J$ categories, was performed for each pair i, j. The number of tests for which the Mantel-Haenszel z exceeds 1.28 (corresponding to a nominal level $p < 0.10$) is tallied, and these particular tests are displayed in detail.

Note that, since $\binom{J}{2}$ statistical tests are performed—one for each pair (i, j)—even if H_0 is true, one would normally expect some significantly positive covariances because of "capitalization on chance." Hence there is a severe multiple inference problem, paralleling a similar problem with the CA condition. Understanding this multiple-inference problem and exploiting possible dependencies among the tests to reduce the problem is an important future step in this research.

The results for the Rasch model are displayed in Table 1 and Table 2. Since LCSN and CSN are equivalent under the Rasch model, H_0 is always true; hence the nine large M-H z's in Table 1 are exclusively due to Type I error (capitalization on chance). Hence Table 1 may be taken as a baseline: if in another model the number of large M-H z's is smaller than for the Rasch model, we may be confident in H_0, and hence (13).

The results for the 2PL simulation are displayed in Tables 3 and 4, and for the 3PL simulation in Tables 5 and 6. Remarkably, fewer large M-H z's were found for these models—where the equivalence of LCSN and CSN is not known theoretically—only two large M-H z's were found for the 2PL simulation, and four were found for the 3PL simulation. Moreover, there are no large M-H z's for the longer tests, suggesting the validity of (12) as $J \to \infty$.

J	# M-H tests	# M-H z's above 1.28
10	45	0
20	190	0
40	780	9
80	3160	0

Table 1: Number of large positive associations for 2000-examinee Rasch simulation (J = test length).

i	j	M-H z	Nominal p	$\hat{\alpha}_{MH}$	$\ln(\hat{\alpha}_{MH})$
32	7	1.403	0.080	1.634	0.491
36	1	1.779	0.038	3.172	1.154
36	8	1.434	0.076	1.640	0.495
37	5	1.497	0.067	2.142	0.762
37	7	2.401	0.008	6.083	1.805
37	27	1.347	0.089	1.295	0.259
38	7	1.502	0.067	2.539	0.932
39	6	1.408	0.080	1.196	0.179
40	34	1.503	0.066	3.343	1.207

Table 2: Details for 40-item Rasch simulation ($\hat{\alpha}_{MH}$ = est. common odds ratio across $\overline{X}_J$ categories).

J	# M-H tests	# M-H z's above 1.28
10	45	1
20	190	1
40	780	0
80	3160	0

Table 3: Number of large positive associations for 2000-examinee 2PL simulation (J = test length).

J	i	j	M-H z	Nominal p	$\hat{\alpha}_{MH}$	$\ln(\hat{\alpha}_{MH})$
10	7	1	1.683	0.046	2.981	1.092
20	18	4	1.460	0.072	2.227	0.801

Table 4: Details for 10- and 20-item 2PL simulations ($\hat{\alpha}_{MH}$ = est. common odds ratio across $\overline{X}_J$ categories).

J	# M-H tests	# M-H z's above 1.28
10	45	0
20	190	4
40	780	0
80	3160	0

Table 5: Number of large positive associations for 2000-examinee 3PL simulation (J = test length).

i	j	M-H z	Nominal p	$\hat{\alpha}_{MH}$	$\ln(\hat{\alpha}_{MH})$
3	1	2.358	0.009	1.570	0.451
4	1	2.435	0.007	1.654	0.503
4	3	1.777	0.038	1.300	0.262
20	1	1.568	0.058	1.411	0.344

Table 6: Details for 20-item 3PL simulation ($\hat{\alpha}_{MH}$ = est. common odds ratio across $\overline{X}_J$ categories).

References

Ash, R. B. (1972). *Real Analysis and Probability.* Academic Press. New York.

Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee's ability*, in Lord, F. M. and Novick, M. R. (1968) *Statistical Theory of Mental Test Scores.* Addison-Wesley. Reading, Massachusetts.

Clarke, B. S. (1990). Private communication.

Clarke, B. S. and Junker, B. W. (1990). Inference from the product of marginals of a dependent likelihood. Working paper.

Cressie, N. and Holland, P. W. (1983). Characterizing the manifest probabilities of latent trait models. *Psychometrika, 48,* 129–141.

Esary, J. D., Proschan, F. and Walkup, D. W. (1967). Association of random variables, with applications. *Annals of Mathematical Statistics, 38* 1466–1474.

Fienberg, S. E. (1986). The Rasch Model, in Kotz, S., Johnson, N. L. and Read, C. B., eds., *Encyclopedia of Statistical Sciences, 7,* 627–632. John Wiley and Sons, Inc. New York.

Grayson, D. A. (1988). Two-group classification in latent trait theory: scores with monotone likelihood ratio. *Psychometrica, 53,* 383–392.

Holland, P. W. (1981). When are item response models consistent with observed data? *Psychometrika, 46,* 79–92.

Holland, P. W. and Rosenbaum, P. R. (1986). Conditional association and unidimensionality in monotone latent trait models. *Annals of Statistics, 14,* 1523–1543.

Jannarone, R. J. (1986). Conjunctive item response theory kernels. *Psychometrika, 51,* 357–373.

Joag-Dev, K. (1983). Independence via uncorrelatedness under certain dependence structures. *Annals of Probability, 11,* 1037–1041.

Joag-Dev, K. and Proschan, F. (1982). Negative association of random variables, with applications. *Annals of Statistics, 10,* 286–295.

Junker, B. W. (1988). Statistical aspects of a new latent trait theory. Ph.D. dissertation. Department of Statistics, University of Illinois at Urbana-Champaign.

Junker, B. W. (1991). Essential independence and likelihood-based ability estimation for polytomous items. To appear, *Psychometrika.*

Lehmann, E. L. (1955). Ordered families of distributions. *Annals of Mathematical Statistics. 26,* 399–419.

Levine, M. V. (1985). Representing ability distributions. Office of Naval Research, Research Report 85-1. Model-based Measurement Laboratory, Department of Educational Psychology, University of Illinois at Urbana-Champaign.

Lord, F. M. (1980). *Application of Item Response Theory to Practical Testing Problems.* Lawrence Erlbaum Associates, Inc. Hillsdale, N. J.

Molenaar, I. (1990). Private communication.

Nandakumar, R. (1987). Refinement of Stout's procedure for assessing latent trait unidimensionality. Ph.D. dissertation. Department of Education, University of Illinois at Urbana-Champaign.

Nandakumar, R. (1989). Validation of Stout's procedure for assessing latent trait unidimensionality with real tests. Paper presented at the 1989 European Meeting of the Psychometric Society, Leuven, Belgium.

Nandakumar, R. and Stout, W. F. (1990). Refinement of a procedure for assessing latent trait unidimensionality. Submitted.

Newman, C. M. and Wright, A. L. (1981). An invariance principle for certain dependent sequences. *Annals of Probability, 9,* 671–675.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* (Expanded edition, 1980.) University of Chicago Press. Chicago, Illinois.

Rosenbaum, P. R. (1984). Testing the conditional independence and monotonicity assumptions of Item Response Theory. *Psychometrika, 49,* 425–436.

Rosenbaum, P. R. (1985). Comparing distributions of item responses for two groups. *British Journal of Mathematical and Statistical Psychology, 38,* 206–215.

Rosenbaum, P. R. (1987). Probability inequalities for latent scales. *British Journal of Mathematical and Statistical Psychology, 40,* 157–168.

Rosenbaum, P. R. (1988). Item bundles. *Psychometrika, 53,* 349–359.

Stout, W. F. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika, 52,* 589–617.

Stout, W. F. (1988). A nonparametric multidimensional IRT approach with applications to ability estimation and test bias. Office of Naval Research, Research Report 88-1. Department of Statistics, University of Illinois at Urbana-Champaign.

Stout, W. F. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika, 55,* 293–325.

Suppes, P. and Zanotti, M. (1981). When are probabilistic explanations possible? *Synthese, 48,* 191–199.

Sympson, J. B. (1987). Applications of a polychotomous IRT model to adaptive mental testing. Paper presented at the Office of Naval Research Contractors' Meeting on Model-based Psychological Measurement, June 1987, University of South Carolina, Columbia, South Carolina.

Zwick, R. (1987) Assessing the dimensionality of NAEP reading data. *Journal of Educational Measurement, 24,* 293–308.

Dr. Terry Ackerman
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. James Algina
1403 Norman Hall
University of Florida
Gainesville, FL 32605

Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

Dr. Ronald Armstrong
Rutgers University
Graduate School of Management
Newark, NJ 07102

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Laura L. Barnes
College of Education
University of Toledo
2801 W. Bancroft Street
Toledo, OH 43606

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

Dr. Isaac Bejar
Law School Admissions
Services
P.O. Box 40
Newtown, PA 18940-0040

Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Ramat Aviv 69978
ISRAEL

Dr. Arthur S. Blaiwes
Code N712
Naval Training Systems Center
Orlando, FL 32813-7100

Dr. Bruce Bloxom
Defense Manpower Data Center
99 Pacific St.
Suite 155A
Monterey, CA 93943-3231

Cdt. Arnold Bohrer
Secue Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningen Astrid
Bruijnstraat
1120 Brussels, BELGIUM

Dr. Robert Breaux
Code 281
Naval Training Systems Center
Orlando, FL 32826-3224

Dr. Robert Brennan
American College Testing
Programs
P. O. Box 168
Iowa City, IA 52243

Dr. Gregory Candell
CTB/McGraw-Hill
2500 Garden Road
Monterey, CA 93940

Dr. John B. Carroll
409 Elliott Rd., North
Chapel Hill, NC 27514

Dr. John M. Carroll
IBM Watson Research Center
User Interface Institute
P.O. Box 704
Yorktown Heights, NY 10598

Dr. Robert M. Carroll
Chief of Naval Operations
OP-01B2
Washington, DC 20350

Dr. Raymond E. Christal
UES LAMP Science Advisor
AFHRL/MOEL
Brooks AFB, TX 78235

Mr. Hua Hua Chung
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
Los Angeles, CA 90089-1061

Director, Manpower Program
Center for Naval Analyses
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Hans F. Crombag
Faculty of Law
University of Limburg
P.O. Box 616
Maastricht
The NETHERLANDS 6200 MD

Ms. Carolyn R. Crone
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Timothy Davey
American College Testing Program
P.O. Box 168
Iowa City, IA 52243

Dr. C. M. Dayton
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Ralph J. DeAyala
Measurement, Statistics,
and Evaluation
Benjamin Bldg., Rm. 4112
University of Maryland
College Park, MD 20742

Dr. Lou DiBello
CERL
University of Illinois
103 South Mathews Avenue
Urbana, IL 61801

Dr. Dattprasad Divgi
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Mr. Hei-Ki Dong
Bell Communications Research
Room PYA-1K207
P.O. Box 1320
Piscataway, NJ 08855-1320

Dr. Fritz Drasgow
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
(2 Copies)

Dr. Stephen Dunbar
224B Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. James A. Earles
Air Force Human Resources Lab
Brooks AFB, TX 78235

Dr. Susan Embretson
University of Kansas
Psychology Department
426 Fraser
Lawrence, KS 66045

Dr. George Englehard, Jr.
Division of Educational Studies
Emory University
210 Fishburne Bldg.
Atlanta, GA 30322

ERIC Facility-Acquisitions
2440 Research Blvd, Suite 550
Rockville, MD 20850-3238

Dr. Benjamin A. Fairbank
Operational Technologies Corp.
5825 Callaghan, Suite 225
San Antonio, TX 78228

Dr. Marshall J. Farr, Consultant
Cognitive & Instructional Sciences
2520 North Vernon Street
Arlington, VA 22207

Dr. P-A. Federico
Code 51
NPRDC
San Diego, CA 92152-6800

Dr. Leonard Feldt
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. Richard L. Ferguson
American College Testing
P.O. Box 168
Iowa City, IA 52243

Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

Dr. Myron Fischl
U.S. Army Headquarters
DAPE-MRR
The Pentagon
Washington, DC 20310-0300

Prof. Donald Fitzgerald
University of New England
Department of Psychology
Armidale, New South Wales 2351
AUSTRALIA

Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Alfred R. Fregly
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Robert D. Gibbons
Illinois State Psychiatric Inst.
Rm 529W
1601 W. Taylor Street
Chicago, IL 60612

Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01003

Dr. Drew Gitomer
Educational Testing Service
Princeton, NJ 08541

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Sherrie Gott
AFHRL/MOMJ
Brooks AFB, TX 78235-5601

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Michael Habon
DORNIER GMBH
P.O. Box 1420
D-7990 Friedrichshafen 1
WEST GERMANY

Prof. Edward Haertel
School of Education
Stanford University
Stanford, CA 94305

Dr. Ronald K. Hambleton
University of Massachusetts
Laboratory of Psychometric
and Evaluative Research
Hills South, Room 152
Amherst, MA 01003

Dr. Delwyn Harnisch
University of Illinois
51 Gerty Drive
Champaign, IL 61820

Dr. Grant Henning
Senior Research Scientist
Division of Measurement
Research and Services
Educational Testing Service
Princeton, NJ 08541

Ms. Rebecca Hetter
Navy Personnel R&D Center
Code 63
San Diego, CA 92152-6800

Dr. Thomas M. Hirsch
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Paul W. Holland
Educational Testing Service, 21-T
Rosedale Road
Princeton, NJ 08541

Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 92010

Ms. Julia S. Hough
Cambridge University Press
40 West 20th Street
New York, NY 10011

Dr. William Howell
Chief Scientist
AFHRL/CA
Brooks AFB, TX 78235-5601

Dr. Lloyd Humphreys
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Steven Hunka
3-104 Educ. N.
University of Alberta
Edmonton, Alberta
CANADA T6G 2G5

Dr. Huynh Huynh
College of Education
Univ. of South Carolina
Columbia, SC 29208

Dr. Robert Jannarone
Elec. and Computer Eng. Dept.
University of South Carolina
Columbia, SC 29208

Dr. Kumar Joag-dev
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright Street
Champaign, IL 61820

Dr. Douglas H. Jones
1280 Woodfern Court
Toms River, NJ 08753

Dr. Brian Junker
Carnegie-Mellon University
Department of Statistics
Schenley Park
Pittsburgh, PA 15213

Dr. Michael Kaplan
Office of Basic Research
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Milton S. Katz
European Science Coordination
Office
U.S. Army Research Institute
Box 65
FPO New York 09510-1500

Prof. John A. Keats
Department of Psychology
University of Newcastle
N.S.W. 2308
AUSTRALIA

Dr. Jwa-keun Kim
Department of Psychology
Middle Tennessee State
University
P.O. Box 522
Murfreesboro, TN 37132

Mr. Soon-Hoon Kim
Computer-based Education
Research Laboratory
University of Illinois
Urbana, IL 61801

Dr. G. Gage Kingsbury
Portland Public Schools
Research and Evaluation Department
501 North Dixon Street
P. O. Box 3107
Portland, OR 97209-3107

Dr. William Koch
Box 7246, Meas. and Eval. Ctr.
University of Texas-Austin
Austin, TX 78703

Dr. Richard J. Koubek
Department of Biomedical
& Human Factors
139 Engineering & Math Bldg.
Wright State University
Dayton, OH 45435

Dr. Leonard Kroeker
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Dr. Jerry Lehnus
Defense Manpower Data Center
Suite 400
1600 Wilson Blvd
Rosslyn, VA 22209

Dr. Thomas Leonard
University of Wisconsin
Department of Statistics
1210 West Dayton Street
Madison, WI 53705

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Charles Lewis
Educational Testing Service
Princeton, NJ 08541-0001

Mr. Rodney Lim
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert L. Linn
Campus Box 249
University of Colorado
Boulder, CO 80309-0249

Dr. Robert Lockman
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

Dr. Richard Luecht
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. George B. Macready
Department of Measurement
Statistics & Evaluation
College of Education
University of Maryland
College Park, MD 20742

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen J. Martin
Office of Chief of Naval
Operations (OP 13 F)
Navy Annex, Room 2832
Washington, DC 20350

Dr. James R. McBride
HumRRO
6430 Elmhurst Drive
San Diego, CA 92120

Dr. Clarence C. McCormick
HQ, USMEPCOM/MEPCT
2500 Green Bay Road
North Chicago, IL 60064

Mr. Christopher McCusker
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Robert McKinley
Educational Testing Service
Princeton, NJ 08541

Mr. Alan Mead
c/o Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Timothy Miller
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152-6800

Ms. Kathleen Moreno
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Headquarters Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Ratna Nandakumar
Educational Studies
Willard Hall, Room 213E
University of Delaware
Newark, DE 19716

Library, NPRDC
Code P201L
San Diego, CA 92152-6800

Librarian
Naval Center for Applied Research
in Artificial Intelligence
Naval Research Laboratory
Code 5510
Washington, DC 20375-5000

Dr. Harold F. O'Neil, Jr.
School of Education - WPH 801
Department of Educational
Psychology & Technology
University of Southern California
Los Angeles, CA 90089-0031

Dr. James B. Olsen
WICAT Systems
1875 South State Street
Orem, UT 84058

Office of Naval Research,
Code 1142CS
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Dr. Judith Orasanu
Basic Research Office
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

Dr. Peter J. Pashley
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Circle, NW
Washington, DC 20036

Dr. James Paulson
Department of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

Dept. of Administrative Sciences
Code 54
Naval Postgraduate School
Monterey, CA 93943-5026

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Malcolm Ree
AFHRL/MOA
Brooks AFB, TX 78235

Mr. Steve Reiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
310B Austin Peay Bldg.
Knoxville, TN 37916-0900

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152-6800

Lowell Schoer
Psychological & Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

Dr. Mary Schratz
4100 Parkside
Carlsbad, CA 92008

Dr. Dan Segall
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Robin Shealy
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaigan
Fujisawa 251
JAPAN

Dr. Randall Shumaker
Naval Research Laboratory
Code 5510
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

Dr. Richard E. Snow
School of Education
Stanford University
Stanford, CA 94305

Dr. Richard C. Sorensen
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Judy Spray
ACT
P.O. Box 168
Iowa City, IA 52243

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Dr. Peter Stoloff
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. William Stout
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

Mr. Brad Sympson
Navy Personnel R&D Center
Code-62
San Diego, CA 92152-6800

Dr. John Tangney
AFOSR/NL, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. Kikumi Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Maurice Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

Mr. Thomas J. Thomas
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Mr. Gary Thomasson
University of Illinois
Educational Psychology
Champaign, IL 61820

Dr. Robert Tsutakawa
University of Missouri
Department of Statistics
222 Math. Sciences Bldg.
Columbia, MO 65211

Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 440
St. Paul, MN 55114

Dr. Frank L. Vicino
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Howard Wainer
Educational Testing Service
Princeton, NJ 08541

Dr. Michael T. Waller
University of Wisconsin-Milwaukee
Educational Psychology Department
Box 413
Milwaukee, WI 53201

Dr. Ming-Mei Wang
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Thomas A. Warm
FAA Academy AAC934D
P.O. Box 25082
Oklahoma City, OK 73125

Dr. Brian Waters
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455-0344

Dr. Ronald A. Weitzman
Box 146
Carmel, CA 93921

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Douglas Wetzel
Code 51
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Rand R. Wilcox
University of Southern
California
Department of Psychology
Los Angeles, CA 90089-1061

German Military Representative
ATTN: Wolfgang Wildgrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washington, DC 20016

Dr. Bruce Williams
Department of Educational
Psychology
University of Illinois
Urbana, IL 61801

Dr. Hilda Wing
Federal Aviation Administration
800 Independence Ave, SW
Washington, DC 20591

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering
Cancer Center
1275 York Avenue
New York, NY 10021

Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
Code 51
San Diego, CA 92152-6800

Dr. Kentaro Yamamoto
02-T
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

Dr. Joseph L. Young
National Science Foundation
Room 320
1800 G Street, N.W.
Washington, DC 20550

Mr. Anthony R. Zara
National Council of State
Boards of Nursing, Inc.
625 North Michigan Avenue
Suite 1544
Chicago, IL 60611